

The Selection of Proper Names in English Dictionaries

CLARENCE L. BARNHART

DICIONARIES ARE THE REPOSITORIES of the vocabulary of the language. As such they attracted the interest and scholarship of Francis Lee Utley, who selected a basic list of the important books in English literature that should be read to provide citations for a college dictionary. Utley also considered the problem of defining proper names in an article, "The Linguistic Component of Onomastics," published in the September, 1963 issue of *Names*. This is a problem that is of obvious practical interest to all dictionary makers.

In varying sizes dictionaries contain selections of the total vocabulary pertaining to a particular segment of the population or to a particular subject or department of knowledge. There are approximately 250,000 words in the working vocabulary of English; college dictionaries, the most common of popular dictionaries, contain from 125,000 to 175,000 terms. Obviously there is a space problem for editors of college dictionaries since the editor must show the facts of spelling, the pronunciation, the inflected forms, the derivatives, the meaning or significance of the word, and the etymology of a little over half of the working vocabulary.

In order to conserve space, facts that may be needed one time out of a thousand by the dictionary user are omitted by the editor—an obsolete or archaic word used only in one context, a technical term used only in advanced study, or a slang term or a dialect term used by comparatively small groups of people for a limited time. Words selected for abridged dictionaries are selected on the basis of frequency, range, and cruciality (e.g., their importance as basic terms in a field of knowledge). Proper names are included in the college dictionaries primarily on the basis of their importance to the prospective user.

Unabridged dictionaries, both historical and descriptive, often contain half a million entries and attempt to be a complete record of the vocabulary for the period covered by the dictionary. Such dictionaries usually do not consider proper names to be a part of the vocabulary to be included. The Preface to Volume I of the *Oxford English Dictionary* (1888 edition, p.vi) states its policy:

. . . it has to be borne in mind, that a Dictionary of the English Language is not a Cyclo-

pædia: the Cyclopædia *describes things*; the Dictionary *explains words*, and deals with the description of things only so far as is necessary in order to fix the exact significations and uses of words. . . . We do not look in a Cyclopædia for the explanation and history of *anon*, *perhaps*, or *busy*; we do not expect in an English Dictionary, information about *Bookbinding*, *Photography*, the *Aniline Dyes*, or the *Bridgewater Treatises*, or mention of *Abyssinia*, *Argynnis*, *Alopecurus*, *Adenia*, or *Blennerheritis*.

In discussing the inconsistency of omitting the proper adjective *African* while including the proper adjective *American*, the *OED* Preface says (p.ix):

. . . the inconsistency is only on the surface; *American* is included, not on its own account, but to help to the better explanation of derived words; and, in every analogous case, it will be found that a proper noun, or adjective thence formed, is included, not for its own sake and as a proper noun, etc., but because it has other uses, or has derivatives for the explanation of which it is of importance. Every such word must, in fact, be looked upon as exceptionally included, and not as forming a precedent for the inclusion of other words of the same class.

The great descriptive dictionary, *The Century Dictionary*, contemporaneous with the *OED*, takes the same attitude toward proper names (Preface to 1911 edition, p. xix):

. . . THE CENTURY DICTIONARY covers to a great extent the field of the ordinary encyclopedia, with this principal difference—that the information given is for the most part distributed under the individual words and phrases with which it is connected, instead of being collected under a few general topics. Proper names, both biographical and geographical, are, of course, omitted except as they appear in derivative adjectives, as *Darwinian* from *Darwin*, or *Indian* from *India*.

A similar but stricter policy of exclusion of proper names guided the editors of *W3* (Preface, p. 6a, col.2):

This dictionary . . . confines itself strictly to generic words and their functions, forms, sounds, and meanings as distinguished from proper names that are not generic.

One looks in vain in these three dictionaries for such entries as *Bucephalus* or *Labiche* or *Kansas* or *Elbe*.

Where does the reader turn for information on *Bucephalus* and similar entries? Usually he is referred to an encyclopedia, and if he follows that advice in this case he will have to persist in order to find the name. *Bucephalus* is not an entry in either the 1960 or 1970 edition of *Collier's Encyclopedia*; it is not entered in the index of either edition, nor is it mentioned under the encyclopedia article on Alexander the Great. *The Encyclopaedia Britannica* has a varied history; the first edition (1771) does not contain *Bucephalus*, but the first American edition does. The 1946 edition lists *Bucephalus* in the index and has a short article in volume 4; the 1966 edition, however, omits it in the encyclopedia proper

and in the index. *The Encyclopedia Americana* (1954) also lists the term in the index and has a somewhat longer article in volume 4 but without a pronunciation; the 1961 edition contains the word with a pronunciation.

In looking up a specific proper name—even one as well known as *Bucephalus*—the reader has a good chance of not finding it at all or of finding it buried in a long article and available only by use of the index. After finding the entry, moreover, the reader will often be unable to discover its pronunciation. Usually none is given in earlier editions of encyclopedias, although there is a tendency to give pronunciations in the newer encyclopedias as they become more dictionary-like in form. But, in general, linguistic facts about proper names—variant spellings, syllabication of the boldface, and pronunciation of the term—are not normally given in encyclopedias. Moreover, encyclopedias are as a rule much more expensive than dictionaries. They tend to be library books, while dictionaries are usually individually owned. This limits the usefulness of the encyclopedia, as it is all too often not available. Many more people use dictionaries than they do encyclopedias.

Giving the linguistic facts about proper names and identifying them is one of the chief functions of college dictionaries, yet unabridged dictionaries as a matter of policy exclude proper names. What is a proper name? How are the proper names accounted for in the grammar of the language? A number of studies and papers have appeared offering a definition of the term *proper name*. In addition to the article by Professor Utley mentioned above, a monograph by John Algeo, entitled *On Defining the Proper Name* and published in 1973 by the University of Florida's Humanities Series, is essential reading for understanding what a proper name is. There are of course those who take the stand that names are not a part of the language. This Algeo regards as a dodge to avoid accounting for facts about language that cannot be conveniently handled by any available linguistic theory.

Algeo examines in detail various criteria for defining proper names that have been suggested: (1) capitalization of proper names versus lower-case words; (2) the syntactic forms of proper names: no plural forms, use without articles, no restrictive modifiers; (3) the uniqueness of proper names in referring to a single individual, and (4) the fact that proper names have no meaning—no characteristics common to several individuals of a group. These four criteria are rejected by Utley, explicitly or implicitly, as well as by Algeo, who by close reasoning and many examples shows their inadequacy.

Algeo does, however, propose a criterion of his own (p. 71): “. . . a proper name is primarily any word X whose meaning can be expressed as ‘entity called X’” Proper nouns differ from common nouns because their meanings are not parallel: a proper name results from an act of

name-giving; a common noun is a name that summarizes essential characteristics of a creature belonging to a class of creatures. Algeo continues, "To know that a creature is appropriately referred to by the word *cat*, it is not necessary to observe anyone calling it 'cat.' But to know that some creature is appropriately called *Pyewacket*, it is necessary to observe some instance of the use of that name with reference to the creature" (p. 71).

In his article, Utley examines the definitions of proper name given by Sir Alan Gardiner, Ernst Pulgram, Otto Jespersen, and Charles F. Hockett, and proceeds to demonstrate how each definition fails *qua* definition by being inductively arrived at or by being wholly dependent on semantic, as against formal, criteria.

Thus Utley rejects the *a priori* criterion of *uniqueness* on the grounds that such a sentence as "Last night there were four Maries here" destroys the uniqueness of *Mary* as a proper name. On the criterion of *determiners* Utley goes only so far as to say that "proper names differ in some fashion from common nouns in their use of the articles *a* and *the* and in their use with adjectives and as adjectives" (p. 168) adding in conclusion, however, that "Pulgram rightly considers the article no final indicator of proper or common noun" (p. 170). Utley also rejects *capitalization* as a criterion, citing many conflicting uses of the capital initial letter. Finally, Utley rejects the criterion that proper names are unlikely to be preceded by adjectives, in view of the use of such common phrases as the *Great* Barnum, the *younger* Dumas, *another* Johnson, a *genuine* Rembrandt, and *amazing* Amsterdam.

If linguistic criteria were adopted as the basis for the exclusion of proper names from a dictionary, dictionary editors would have to exclude the *Cenozoic* (geological age), the *U.N.*, the *Mafia*, the *Devil*, and *earth* (the planet) since they satisfy most if not all grammatical criteria for properness. Such standards are difficult to apply: witness *W3*'s entry of the planet *earth* (def. 5 *often cap*) but its exclusion of the planet *Neptune*. Such entries as *Age of Mammals* (the Cenozoic), *Aircav* (a unit of the U.S. armed forces), *Amtrak* (a system of rail passenger service), *BASIC* (a computer language), *Big Bang* (a cosmic explosion 10 to 15 billion years ago), *Bennett* (a comet sighted in 1969), and *art deco* (a style of design) would have to be omitted as proper names. On the other hand such capitalized terms as *Uncle Tom*, *White Paper*, *Chicano*, *May Day*, and *Rhodes scholar* would be admitted as common nouns. The truth of the matter is that both types of terms belong in a popular dictionary: they have currency in the literature of the day; most of them would not be entered in any other reference work. Frequency, range, and cruciality are the criteria for including or excluding entries in a dictionary; grammatical criteria do not work.

The inclusion of proper names is also helpful in the explanation of names and adjectives derived from them. The proper names are located

in space or time and important linguistic facts such as pronunciation and variant spellings or names are given. The definition of *Ishmael*, the secondary sense of which is "an outcast, a person like Ishmael" can be helped by the definition of the primary term. So can *Shelleyan* be helped by the entry of *Shelley*.

The primary terms are so important that their attempted omission results in outlandish treatment that can only puzzle and discourage the average dictionary user. *W3*, which has been the most consistent dictionary in excluding proper names, often puts a definition of the proper name under the proper adjective, thus playing hide and seek with the user of the dictionary. *Labrador*, the peninsula in Canada, is not entered in the dictionary but is explained in the etymology for *Labrador* (the adjective) as follows: "[fr. *Labrador*, peninsula, Newfoundland and Quebec, provinces in Canada]" and again under *Labradorean*: "[*Labrador*, peninsula, Canada + *E -an*]." The *OED*, which also had the policy of excluding proper names, enters *Labrador* as a main entry and defines it as "the name of a large peninsula in British North America," including it of course as necessary to the understanding of such collocations as *Labrador blue*, *Labrador duck*, and so on. *The Century* does not enter *Labrador* and does not identify *Labrador* in the etymology of *Labradorian*: "[*Labrador* + *-ian*]." Each dictionary has its facts selected and arranged in a different way: the *OED* puts the primary sense first, *W3* puts the primary sense in the etymology, and the *Century* puts *Labrador* in the etymology without explanation. There are justifications for each decision, but most American editors of college dictionaries today would follow the policy of entering and identifying *Labrador* as a more important (frequent) term than *Labradorean*. *Labradorean* indeed may be omitted or treated as a run-on entry. On linguistic grounds this may be the wrong policy, but the fact is that the user of the dictionary is usually seeking information about *Labrador* and not *Labradorean*. If he does seek information about *Labradorean*, his curiosity is gratified by simply connecting *Labradorean* to *Labrador*.

In sum, the function of a popular dictionary is to give desired and reliable linguistic information about specific words—whether or not they are proper names; the function of an encyclopedia is to deal with whole fields of knowledge. A descriptive dictionary that selects entries on linguistic principles may be satisfying to some linguists, but it is bound to be inadequate as a source of the information needed by the educated user of dictionaries today. If strict linguistic criteria were to be adopted in the selection of terms to be included in dictionaries, many current words and terms would have to be omitted on theoretical grounds, and information about many terms, especially those that are used to identify particular or individual objects, would be unavailable.